Efficient Sharing of Resources in Distributed Systems

Debessay (Debish) Fesehaye Kassa

Dept. of Computer Science
University of Illinois at Urbana-Champaign

May 19, 2013

Max/Min Fairness Using Efficient Sharing (ES): Method 1

The rate

$$Rd, u(t) = \frac{C_{d,u} - \frac{q_{d,u}(t-\tau)}{\tau}}{\hat{N}_{d,u}(t-\tau)}$$

Max/Min: *Fractional flows*

$$\hat{N}_{d,u}(t-\tau) = \frac{S_{d,u}(t)}{R_{d,u}(t-\tau)} = \sum_j^{N_{d,u}(t-\tau)} \frac{R^j_{d,u}(t)}{R_{d,u}(t-\tau)}$$

QoS, SLA check

$$S_{d,u}(t) = \sum_j^{N_{d,u}(t-\tau)} \wp^j_{d,u} R^j_{d,u}(t)$$

Per flow weighted share at the link

$$\wp^j_{d,u} R_{d,u}(t)$$

Notations:

- $C_{d,u}$ = Link capacity for down (d) and up (u) links
- $q_{d,u}(t)$ = Queue size
- τ = Control interval
- $N_{d,u}(t)$ = Number of flows
- $\wp^j_{d,u} = \frac{R^j_{d,u}(t+\tau)}{R^j_{d,u}(t)} =$ Priority weight of flow j
- $R^j_{d,u}(t)$ = Rate of flow j

Efficient Sharing Example

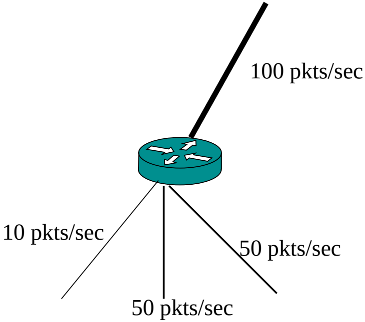
MaxMin fairness example

- $C_u = 100 pkts/sec$
- $q_u(t) = 0 pkts$
- $\tau = 1 sec$
- $N_u = 3$
- $R_u^1(t) = 10, R_u^2 = 50, R_U^3 = 50$ pkts/sec
- $\wp_1^u = \wp_2^u = \wp_3^u = 1$
- $R_u(t - \tau) = 50$ pkts/sec

Efficient Sharing Example

MaxMin fairness example

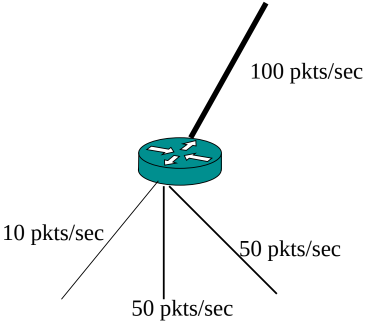
- $C_u = 100 pkts/sec$
- $q_u(t) = 0pkts$
- $\tau = 1sec$
- $N_u = 3$
- $R_u^1(t) = 10, R_u^2 = 50, R_U^3 = 50$ pkts/sec
- $\wp_1^u = \wp_2^u = \wp_3^u = 1$
- $R_u(t - \tau) = 50$ pkts/sec



Efficient Sharing Example

MaxMin fairness example

- $C_u = 100 pkts/sec$
- $q_u(t) = 0 pkts$
- $\tau = 1 sec$
- $N_u = 3$
- $R_u^1(t) = 10, R_u^2 = 50, R_U^3 = 50$ pkts/sec
- $\wp_1^u = \wp_2^u = \wp_3^u = 1$
- $R_u(t - \tau) = 50$ pkts/sec



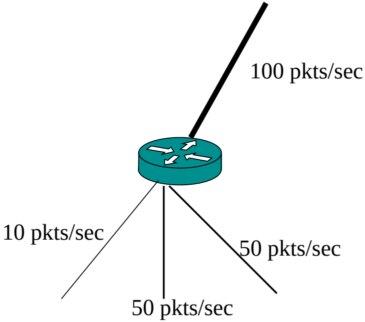
Then ES rate is

$$R(t) = \frac{100}{\frac{10}{50} + \frac{50}{50} + \frac{50}{50}} = 45.45 \text{ pkts/sec.}$$

Efficient Sharing Example

MaxMin fairness example

- $C_u = 100pkts/sec$
- $q_u(t) = 0pkts$
- $\tau = 1sec$
- $N_u = 3$
- $R_u^1(t) = 10, R_u^2 = 50, R_U^3 = 50$ pkts/sec
- $\wp_1^u = \wp_2^u = \wp_3^u = 1$
- $R_u(t-\tau) = 50$ pkts/sec



Then ES rate is

$$R(t) = \frac{100}{\frac{10}{50}+\frac{50}{50}+\frac{50}{50}} = 45.45 \text{ pkts/sec}.$$

Processor sharing (PS):

$$R(t) = \frac{100}{3} = 33.33 \text{ pkts/sec}.$$

Max/Min Fairness Using ES: Method 2

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- In this method we set $R(t-\tau) = \frac{X}{N}$, which is the processor sharing rate.
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq R(t-\tau)$ as a source j cannot send higher than its bottleneck fair share.
- Then the ES rate, $R(t)$ can be given as $R(t) = \frac{X}{\sum_j^N \left(\frac{R^j(t)}{R(t-\tau)}\right)} = \frac{X^2}{N \sum_j^N R^j(t)}$.

Max/Min Fairness Using ES: Method 2

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- In this method we set $R(t-\tau) = \frac{X}{N}$, which is the processor sharing rate.
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq R(t-\tau)$ as a source j cannot send higher than its bottleneck fair share.
- Then the ES rate, $R(t)$ can be given as $R(t) = \frac{X}{\sum_j^N (\frac{R^j(t)}{R(t-\tau)})} = \frac{X^2}{N\sum_j^N R^j(t)}$.

Then the rate for the example above, in the first iteration (round) becomes

$$R(t) = \frac{100}{\frac{10}{33.33} + \frac{33.33}{33.33} + \frac{33.33}{33.33}} = 43.48 \text{ pkts/sec.}$$

In the second iteration (round)

$$R(t) = \frac{100}{\frac{10}{43.48} + \frac{43.48}{43.48} + \frac{43.48}{43.48}} = 44.84 \text{ pkts/sec.}$$

Values of ES rate in next iterations (rounds or RTTs)

44.9842555105713, 44.9984250551231, 44.9998425005512, 44.9999842500055,

Generalized Efficient Sharing (GES)

With priority weight, $\wp^j$ of flow j, the GES rate $R(t)$ can be given as $R(t) = \frac{X}{\sum_j^N (\frac{\wp^j R^j(t)}{R(t-\tau)})} = \frac{X^2}{N\sum_j^N \wp^j R^j(t)}$. Then source j's weighted share becomes $\wp^j R(t)$.

Max/Min Fairness Using ES: Method 3

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq \frac{C}{N(t-\tau)}$ as a source j cannot send higher than its bottleneck fair share.
- In this method, first

  $N(t) \leftarrow 0$; $\hat{N}(t) \leftarrow 0$; $\hat{X}(t) \leftarrow 0$;
  for each flow j **do**
    $N(t) \leftarrow N(t) + 1$;
    if $(R^j(t) < \frac{X}{N(t-\tau)}$ **then**
      $\hat{X}(t) \leftarrow \hat{X}(t) + R^j(t)$;
      $\hat{N}(t) \leftarrow \hat{N}(t) + 1$;
    end if
  end for

- The ES rate $R(t)$ is then given by $R(t) = \frac{X - \hat{X}(t)}{\tilde{N}(t)}$; $\tilde{N}(t) = N(t) - \hat{N}(t)$.

Max/Min Fairness Using ES: Method 3

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq \frac{C}{N(t-\tau)}$ as a source j cannot send higher than its bottleneck fair share.
- In this method, first

  $N(t) \leftarrow 0$; $\hat{N}(t) \leftarrow 0$; $\hat{X}(t) \leftarrow 0$;
  for each flow j **do**
    $N(t) \leftarrow N(t) + 1$;
    if $(R^j(t) < \frac{X}{N(t-\tau)}$ **then**
      $\hat{X}(t) \leftarrow \hat{X}(t) + R^j(t)$;
      $\hat{N}(t) \leftarrow \hat{N}(t) + 1$;
    end if
  end for
- The ES rate $R(t)$ is then given by $R(t) = \frac{X - \hat{X}(t)}{\tilde{N}(t)}$; $\tilde{N}(t) = N(t) - \hat{N}(t)$.

$\hat{N}(t)$ and $\hat{X}(t)$ are number of flows and sum of rates of the flows bottlenecked at other resources.

Max/Min Fairness Using ES: Method 3

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq \frac{C}{N(t-\tau)}$ as a source j cannot send higher than its bottleneck fair share.
- In this method, first

  $N(t) \leftarrow 0$; $\hat{N}(t) \leftarrow 0$; $\hat{X}(t) \leftarrow 0$;
  for each flow j **do**
  $\quad N(t) \leftarrow N(t) + 1$;
  $\quad$ **if** $(R^j(t) < \frac{X}{N(t-\tau)}$ **then**
  $\quad\quad \hat{X}(t) \leftarrow \hat{X}(t) + R^j(t)$;
  $\quad\quad \hat{N}(t) \leftarrow \hat{N}(t) + 1$;
  $\quad$ **end if**
  end for
- The ES rate $R(t)$ is then given by $R(t) = \frac{X - \hat{X}(t)}{\tilde{N}(t)}$; $\tilde{N}(t) = N(t) - \hat{N}(t)$.

$\hat{N}(t)$ and $\hat{X}(t)$ are number of flows and sum of rates of the flows bottlenecked at other resources.

Then rate for the example above

$$R(t) = \frac{100-10}{2} = 45.0 \text{ pkts/sec}.$$

Max/Min Fairness Using ES: Method 3

MaxMin fairness: Efficient Sharing

- Given a resource with capacity X units/sec to be shared by N sources,
- Each source's bottleneck fair (ES) share rate is denoted with $R^j(t)$.
- We also have $R^j(t) \leq \frac{C}{N(t-\tau)}$ as a source j cannot send higher than its bottleneck fair share.
- In this method, first

  $N(t) \leftarrow 0$; $\hat{N}(t) \leftarrow 0$; $\hat{X}(t) \leftarrow 0$;
  for each flow j **do**
  $\quad N(t) \leftarrow N(t) + 1$;
  $\quad$**if** $(R^j(t) < \frac{X}{N(t-\tau)}$ **then**
  $\quad\quad \hat{X}(t) \leftarrow \hat{X}(t) + R^j(t)$;
  $\quad\quad \hat{N}(t) \leftarrow \hat{N}(t) + 1$;
  $\quad$**end if**
  end for
- The ES rate $R(t)$ is then given by $R(t) = \frac{X - \hat{X}(t)}{\tilde{N}(t)}$; $\tilde{N}(t) = N(t) - \hat{N}(t)$.

$\hat{N}(t)$ and $\hat{X}(t)$ are number of flows and sum of rates of the flows bottlenecked at other resources.

Then rate for the example above

$$R(t) = \frac{100-10}{2} = 45.0 \text{ pkts/sec.}$$

Generalized Efficient Sharing (GES)

With priority weight, $\wp^j$ of flow j, the GES rate $R(t)$ can be given as

$R(t) = \frac{X - \hat{X}(t)}{\sum_j^{\tilde{N}(t)} \wp^j}$, where the sum is over the $\tilde{N}(t)$ flows bottlnecked at the current resource with capacity X units/sec.

Then source j's weighted share at the current resource becomes $\wp^j R(t)$.

Cross-Layer Approach for General Networks

Problems

Cross-Layer Approach for General Networks

Problems

Which path and at what rate?

Cross-Layer Approach for General Networks

Problems

Which path and at what rate?

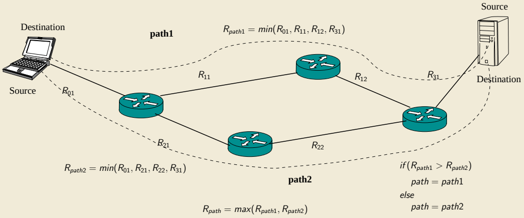
Solution

Cross-Layer Approach for General Networks

Problems

Which path and at what rate?

Solution

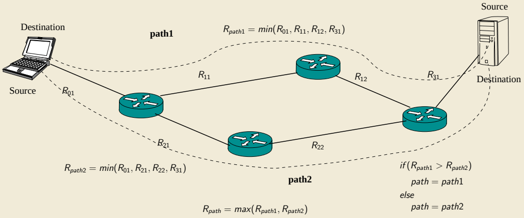


Cross-Layer Approach for General Networks

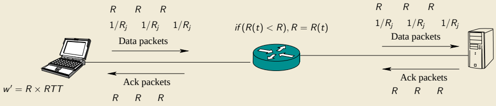
Problems

Which path and at what rate?

Solution



R_{path} from data to ack packets.



Notations:

- R = bottleneck link rate. Source sets it to its desired rate
- R_j = current source sending rate
- $R_{nm} = R(t)$ = fair rate calculated by each router every control interval τ
- w' = flows new cwnd

Name

We call our cross-layer approach for general networks a *Quick Congestion control Protocol (QCP)*

Max/Min Rate for General Networks: Method 1

Notations

- $C =$ Link capacity
- $q(t) =$ Queue size
- $\tau =$ Control interval
- $L =$ Number of all packets
- $R_j =$ Flow rate at packet j
- $N =$ Number of flows
- $R_i =$ Sending rate of flow i
- $n_i =$ Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; $R_j^T =$ Target rate

Max/Min Rate for General Networks: Method 1

Notations

- $C =$ Link capacity
- $q(t) =$ Queue size
- $\tau =$ Control interval
- $L =$ Number of all packets
- $R_j =$ Flow rate at packet j
- $N =$ Number of flows
- $R_i =$ Sending rate of flow i
- $n_i =$ Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; $R_j^T =$ Target rate

The rate

$$R(t) = \frac{C - q(t)/\tau}{\frac{1}{\tau}\sum_{j=1}^{L}\frac{1}{R_j}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{n_i/\tau_i}{R_i}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{r_i}{R_i}}.$$

Comment

- No per flow state
- Just the denominator sum
- Rate updated every τ
- Values reset every τ

Max/Min Rate for General Networks: Method 1

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- R_j = Flow rate at packet j
- N = Number of flows
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Comment

- No per flow state
- Just the denominator sum
- Rate updated every τ
- Values reset every τ

The rate

$$R(t) = \frac{C - q(t)/\tau}{\frac{1}{\tau}\sum_{j=1}^{L}\frac{1}{R_j}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{n_i/\tau_i}{R_i}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{r_i}{R_i}}.$$

QoS

$$R(t) = \frac{Cd - q(t)}{\sum_{j=1}^{L}(\omega_j/R_j)};\ R_j = \omega_j R(t) \text{ at source.}$$

Max/Min Rate for General Networks: Method 1

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- R_j = Flow rate at packet j
- N = Number of flows
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Comment

- No per flow state
- Just the denominator sum
- Rate updated every τ
- Values reset every τ

The rate

$$R(t) = \frac{C - q(t)/\tau}{\frac{1}{\tau}\sum_{j=1}^{L}\frac{1}{R_j}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{n_i/\tau_i}{R_i}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{r_i}{R_i}}.$$

QoS

$$R(t) = \frac{Cd - q(t)}{\sum_{j=1}^{L}(\omega_j/R_j)};\ R_j = \omega_j R(t) \text{ at source.}$$

SLA Violation

If $\sum_j^L R_j > LR(t)$

Max/Min Rate for General Networks: Method 1

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- R_j = Flow rate at packet j
- N = Number of flows
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Comment

- No per flow state
- Just the denominator sum
- Rate updated every τ
- Values reset every τ

The rate

$$R(t) = \frac{C - q(t)/\tau}{\frac{1}{\tau}\sum_{j=1}^{L}\frac{1}{R_j}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{n_i/\tau_i}{R_i}} = \frac{C - q(t)/\tau}{\sum_{i=1}^{N}\frac{r_i}{R_i}}.$$

QoS

$$R(t) = \frac{Cd - q(t)}{\sum_{j=1}^{L}(\omega_j/R_j)};\ R_j = \omega_j R(t) \text{ at source.}$$

SLA Violation

If $\sum_j^L R_j > LR(t)$

Max/min fairness

Fractional flows: $R_j := max(R_j, R(t-\tau))$

Max/Min Rate Example for General Networks (Method 1)

MaxMin Fairness Example

- $C = 100 pkts/sec$
- $q(t) = 0 pkts$
- $\tau = 1 sec$
- $N = 3$
- $R_1 = 10, R_2 = 50, R_3 = 70$ pkts/sec
- $n_1 = 10, n_2 = 50, n_3 = 70$ pkts
- $L = n_1 + n_2 + n_3 = 130$ pkts
- $R(t - \tau) = 40$ pkts/sec
- $\Rightarrow R_1 := max(R_1, R(t - \tau)) = 40$ pkts/sec

Max/Min Rate Example for General Networks (Method 1)

MaxMin Fairness Example

- $C = 100 pkts/sec$
- $q(t) = 0 pkts$
- $\tau = 1 sec$
- $N = 3$
- $R_1 = 10, R_2 = 50, R_3 = 70$ pkts/sec
- $n_1 = 10, n_2 = 50, n_3 = 70$ pkts
- $L = n_1 + n_2 + n_3 = 130$ pkts
- $R(t - \tau) = 40$ pkts/sec
- $\Rightarrow R_1 := max(R_1, R(t - \tau)) = 40$ pkts/sec

Then

$$R(t) = \frac{100}{\frac{10}{40} + \frac{50}{50} + \frac{70}{70}} = 44.44 \text{ pkts/sec.}$$

Processor sharing:

$$R(t) = \frac{100}{3} = 33.33 \text{ pkts/sec.}$$

Max/Min Rate for General Networks: Method 3

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- $\hat{L}$ = Number of all packets bottlenecked at other links (resources)
- R_j = Flow rate at packet j
- $N(t)$ = Number of flows at time t
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Max/Min Rate for General Networks: Method 3

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- $\hat{L}$ = Number of all packets bottlenecked at other links (resources)
- R_j = Flow rate at packet j
- $N(t)$ = Number of flows at time t
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Every control interval τ

$L \leftarrow 0$; $\hat{L}(t) \leftarrow 0$; $\ell \leftarrow 0$; $\hat{\ell} \leftarrow 0$;
for each arriving packet j **do**
  $L \leftarrow L + 1$;
  $\ell \leftarrow \ell + \frac{1}{R_j}$;
  if $(R_j < R(t-\tau)$ **then**
    $\hat{L} \leftarrow \hat{L} + 1$;
    $\hat{\ell} \leftarrow \hat{\ell} + \frac{1}{R_j}$;
  end if
end for

The max/min ($R^M(t)$) and equal $R(t)$ rates are

$$R^M(t) = \frac{C - \frac{q(t)}{\tau} - \frac{\hat{L}}{\tau}}{\frac{1}{\tau}(\ell - \hat{\ell})}, \text{ and } R(t) = \frac{C - \frac{q(t)}{\tau}}{\frac{1}{\tau}\ell}.$$

Max/Min Rate for General Networks: Method 3

Notations

- C = Link capacity
- $q(t)$ = Queue size
- τ = Control interval
- L = Number of all packets
- $\hat{L}$ = Number of all packets bottlenecked at other links (resources)
- R_j = Flow rate at packet j
- $N(t)$ = Number of flows at time t
- R_i = Sending rate of flow i
- n_i = Num of pkts of flow i
- $r_i = \frac{n_i}{\tau}$ Actual arrival rate
- $\omega_j = \frac{R_j^T}{\hat{R}_j}$; R_j^T = Target rate

Every control interval τ

$L \leftarrow 0$; $\hat{L}(t) \leftarrow 0$; $\ell \leftarrow 0$; $\hat{\ell} \leftarrow 0$;
for each arriving packet j **do**
  $L \leftarrow L + 1$;
  $\ell \leftarrow \ell + \frac{1}{R_j}$;
  if $(R_j < R(t-\tau)$ **then**
    $\hat{L} \leftarrow \hat{L} + 1$;
    $\hat{\ell} \leftarrow \hat{\ell} + \frac{1}{R_j}$;
  end if
end for

The max/min $(R^M(t))$ and equal $R(t)$ rates are

$$R^M(t) = \frac{C - \frac{q(t)}{\tau} - \frac{\hat{L}}{\tau}}{\frac{1}{\tau}(\ell - \hat{\ell})}, \text{ and } R(t) = \frac{C - \frac{q(t)}{\tau}}{\frac{1}{\tau}\ell}.$$

QoS

$R^M(t) = \frac{C - \frac{q(t)}{\tau} - \frac{\hat{L}}{\tau}}{\sum_{j=1}^{\tilde{L}}(\omega_j / R_j)}$, $\tilde{L} = L - \hat{L}$; $R_j = \omega_j R^M(t)$ at source. Sum over $\tilde{L}$ packets of locally bottlnecked flows